Monthly load workflow Jamie McGarty

Preliminary tasks   [edit]

The working directory in which all the following takes place is:

> G:\TechServ\ESM\eresources cataloging\SerialsSolutions ebook MARC processing\

Make sure HSL and Law have finished loading their records from the previous month   [edit]

- HSL: This directory should be empty: H:\HSL Acquisitions\SerSol Ebook MARC
- Law: record note here when Jesse is finished loading: no December 2016 records for Law.

Download and extract monograph MARC records   [edit]

- Download from: ftp://[user]:[pw]@ftp.serialssolutions.com/
- Extract monograph .mrc files
- Move .mrc files to: data/ssmrc/orig/YYYY

We don't retain the original .zip files or any of the included spreadsheets

Run monthly processing script   [edit]

For routine use, this:

- renames the SerialsSolutions MARC files to a format easier to identify/access from within the scripts
- change the ssj/ssib 001 prefixes so we can keep SerialsSolutions ebook MARC separate from ejournal MARC in our system
- make sure all packages represented in incoming MARC are accounted for in the pkg_data.csv file

Each of these functions is run manually, in the order specified below. The other functions are in development/experimental or are very rarely used.

In Powershell:

> ruby bin/monthly_processing.rb

- Enter YYYYMM when prompted
- Rename SerSol files
  - Uses standard format including YYYYMM01 date format and type of record (add/change/delete)
  - Appends "o" to indicate these are the original files
- Localize ssj numbers in files
  - Creates a copy of each original file for the month same name without 'o' on the end.
  - Edits 001 in new file, changing ssj to sse, ssib to sseb
- Check packages in add file
  - Looks for records in packages not tracked in data/pkg_list.csv Important step for (1) loading only what we want to; and (2) adding appropriate 773 and 506 to records
  - Will report name changes as new packages
  - If there are no new packages found, quit out of monthly_processing.rb. If there ARE new packages, see following section.

Handling new packages in MARC delivery   [edit]

Check SerialsSolutions Client Center Knowledgebase notifications page   to see if it is a database/package title change. If so:

- Add new name to beginning of name column in data/pkg_list.csv separate this new value from existing value(s) with: ;;; save and close data/pkg_list.csv
- Move prior package name to *AltTitle* column of master spreadsheet. Change *Title* column to reflect new name.

Otherwise:

- Consult with Kristina Spurgin

Get current SerSol ebook data out of Millennium   [edit]

- Use a big enough list. On 20161017, there were 91032 records.
- Run create list query for bib records index query index = o begin = sse0 end = sseb99999999999999
- Export results: fields
  - b, 001
  - b, location (26)
  - Field delimiter <9>
  - Text qualifier <none>
  - Browse: {working directory}\data\mill_data.txt
- Open resulting .txt file in Excel
- Save file as data/mill_data.csv

Run bin/splitter.rb   [edit]

- Adds 773 and 506 fields per package as specified in data/pkg_data.csv
- Does some basic automated cleanup of known issues
- Sorts through all provided records and splits into the following files, placed in data/ssmrc/split_lib:
  - YYYYMM01_aal_add.mrc New records for AAL. All records in the adds file that are in packages we load, and all records from the changes file that are newly in packages we load (and are not already in the catalog). 773 fields are added for stats counting.
  - YYYYMM01_aal_delete.mrc Delete records for AAL. All records in the deletes file that are loaded in the catalog and have er,eb as the location.
  - YYYYMM01_change.mrc Change records for everyone, to be loaded by AAL. 773 fields are added to AAL package records for stats counting.
  - YYYYMM01_hsl_add.mrc New records for HSL. All records in the adds file that are in packages HSL tracks/loads, and all records from the changes file that are newly in packages HSL loads (and are not already in the catalog).
    - HSL prefers to get their adds split up by package, so there will be some number of files with the following file naming convention: YYYYMM01_hsl_add_PACKAGENAME.mrc
  - YYYYMM01_hsl_delete.mrc Delete records for HSL. All records in the deletes file that are loaded in the catalog and have noh in the location
  - YYYYMM01_law_add.mrc New records for LAW. All records in the adds file that are in packages LAW tracks/loads, and all records from the changes file that are newly in packages LAW loads (and are not already in the catalog).
  - YYYYMM01_law_delete.mrc Delete records for LAW. All records in the deletes file that are loaded in the catalog and have k in the location

Send LAW their Add and Delete records [edit]

- If there are no records for LAW:
  - Email (as eres_cat) Jesse Griffin and let him know
  - Update note above that there were no records sent to Law
- If there ARE records...
  - Send (as eres_cat) to Jesse Griffin via email (as attachments)
  - Jesse is supposed to notify us when he's done processing/loading these, at which point we'd update the note above that Law is done for the month

Pass on HSL files [edit]

Put the following files in the directory: H:\HSL Acquisitions\SerSol Ebook MARC

- All files with pattern: YYYYMM01_hsl_add_PACKAGENAME.mrc
- YYYYMM01_hsl_delete.mrc

Email Dee Stuckey and Rebecca Butler (as eres_cat) with the location of the files, letting them know the files are ready.

Delegate add and change file editing to Connie Israel [edit]

- Create a task in eres_cat shared mailbox tasks
- Task subject: prep MARC file: SerialsSolutions Adds and Changes MONTH
- Task category: CI TASK
- Task body:

```
Adds:
File: PATH-TO-MRC-FILE

Changes:
File: PATH-TO-MRC-FILE

Instructions: https://intranet.lib.unc.edu/wikis/staff/index.php/Serials_Solutions_360_MARC_for_ebooks_implementation#Monthly_load_workflow_-_E-Resources_Cataloging_Assist
```

Connie Israel will send "to load" emails with load notes to eres_cat@unc.edu when this work is complete.

Prep and load delete file [edit]

Prep [edit]

- Open/break in MARCedit: YYYYMM01_aal_delete.mrc
  - Do not use translate to UTF8 option
- Run MARCedit task *sdel SerSol delete prep* (ShiftF7)
  - *Deletes everything except the 001 for overlay. Allows faster loading, since whole records do not have to be replaced. Keeps spurious dupes from being reported by Headings Report.*
- Compile as .mrc in data/ssmrc/split_load

Load [edit]

- Load with Load Table 9 (Sersol Ebook file)
- After loading:
  - use Global Update: suppress_bibs
  - pull attached items into a list and Global Update: withdraw_item
  - Record as a batch ebook withdrawal in your stats
  - Request batch record deletion

When all 3 files have been loaded: Create an Outlook Task to notify Connie Israel that monthly process is complete [edit]

Subject: SerialsSolutions ebooks post load checks and edits

Category: CI TASK

Due date: next business day

Body of task:

The monthly SerialsSolutions ebooks processing and load for MONTH is complete.

Please take care of any postload checks/edits on individual records listed here:https://intranet.lib.unc.edu/wikis/staff/index.php/SerialsSolutions_monograph_MARC__post load_checks_on_individual_records_done_by_EResources_Cataloging

Clean up files [edit]

Orig directory

- Delete localized .mrc files (i.e. those that don't have 'o' added to file name)
- Remove "o" from original .mrc file names

split load directory

- Delete the previous month's files

split lib directory

- Delete any .mrks
- Zip remaining .mrc files and name YYYYMM.zip
- Delete previous month's .zip file